

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

1 CCTGCCTGGT CCTCTGTGCC TGGTGGGGTG GGGGTGCCAG GTGTGTCCAG  
51 AGGAGCCCAT TTGGTAGTGA GGCAGGTATG GGGCTAGAAG CACTGGTGCC  
101 CCTGGCCGTG ATAGTGGCCA TCTTCCTGCT CCTGGTGGAC CTGATGCACC  
151 GGCGCCAACG CTGGGCTGCA CGCTACTCAC CAGGCCCCCT GCCACTGCCC  
201 GGGCTGGGCA ACCTGTGCA TGTGGACTTC CAGAACACAC CATACTGCTT  
251 CGACCAAGTT CGGCGCCGCT TCGGGGACGT GTTCAGCCTG CAGCTGGCCT  
301 GGACGCCGTG GGTCTGTGCT AATGGGTGG CGGCCGTGCG CGAGGCGCTG  
351 GTGACCCACG GCGAGGACAC CGCCGACCG CCGCCTGTGC CCATCACCCA  
401 GATCCTGGGT TTTGGGCCGC GTTCCCAAGG ACGCCCCTTT CGCCCCAACG  
451 GTCTCTTGA CAAAGCCGTG AGCAACGTGA TCGCCTCCCT CACCTGCGGG  
501 CGCCGCTTCG AGTACGACGA CCCTCGCTTC CTCAGGCTGC TGGACCTAGC  
551 TCAGGAGGGA CTGAAGGAGG AGTCGGGCTT TCTGCGCGAG GTGCTGAATG  
601 CTGTCCCGCT CCTCTGTCAT ATCCCAGCGC TGGCTGGCAA GGTCTACGC  
651 TTCCAAAGG CTTTCCTGAC CCAGCTGGAT GAGCTGCTAA CTGAGCACAG  
701 GATGACCTGG GACCCAGCCC AGCCCCCCC AGACCTGACT GAGGCCTTCC  
751 TGGCAGAGAT GCGAAGGCC AAGGGAACC CTGAGAGCAG CTTCAATGAT  
801 GAGAACCTGC GCATAGTGGT GGCTGACCTG TTCTCTGCCG GGATGGTGAC  
851 CACCTCGACC ACGCTGGCCT GGGGCTCCT GCTCATGATC CTACATCCGG  
901 ATGTGCAGCG CCGTGTCCTA CAGGAGATCG ACGACGTGAT AGGGCAGGTG  
951 CGGCGACGAG AGATGGGTGA CCAGGCTCAC ATGCCCTACA CCACTGCCGT  
1001 GATTCTAGAG GTGCAGCGCT TTGGGGACAT CGTCCCCCTG GGTGTGACCC  
1051 ATATGACATC CCGTGACATC GAAGTACAGG GCTTCCGCAT CCTAAGGGA  
1101 ACGACACTCA TCACCAACCT GTCATCGGTG CTGAAGGATG AGGCCGTCTG  
1151 GGAGAAGCCG TTCCGCTTCC ACCCGAACA CTTCTGGAT GCCCAGGGC  
1201 ACTTTGTGAA GCCGGAGGCC TTCCTGCCTT TCTCAGCAGG CCGCCGTGCA  
1251 TGCTCGGGG AGCCCCCTGGC CCGCATGGAG CTCTTCCTCT TCTTCACCTC  
1301 CCTGCTGCAG CACTTCAGCT TCTCGGTGCC CACTGGACAG CCCCAGGCCA  
1351 GGCACCATGG TGTCTTTGCT TTCTTGGTGA CCCCATCCCC CTATGAGCTT  
1401 TGTGCTGTGC CCCGCTAGAA TGGGTACCT AGTCCCCAGC CTGCTCCCTA  
1451 GCCAGAGGCT CTAATGTACA ATAAAGCAAT GTGGTAGTTC CAAAAA  
1501 AAAAAA AAAAAA AAAAAA AAAAAA  
(SEQ ID NO: 1)

#### FEATURES:

5'UTR: 1 - 77  
Start Codon: 78  
Stop Codon: 1416  
3'UTR: 1419

#### Homologous proteins:

##### Top 10 BLAST Hits

	Score	E
CRA 18000004889269 /altid=gi 181304 /def=gb AAA53500.1  (M33388...	884	0.0
CRA 18000004927597 /altid=gi 4503223 /def=ref NP_000097.1  cyto...	883	0.0
CRA 18000004923926 /altid=gi 181306 /def=gb AAA35737.1  (M33189...	864	0.0
CRA 18000005007118 /altid=gi 2493367 /def=sp Q29488 CPDH_MACFA ...	827	0.0
CRA 18000005100319 /altid=gi 3913340 /def=sp O18992 CPDJ_CALJA ...	800	0.0
CRA 18000004884804 /altid=gi 486997 /def=pir  S37284 cytochrome...	682	0.0
CRA 18000004889271 /altid=gi 522195 /def=gb AAA36403.1  (M24499...	673	0.0
CRA 18000004884803 /altid=gi 461826 /def=sp Q01361 CPDE_BOVIN C...	669	0.0
CRA 18000004939934 /altid=gi 117244 /def=sp P13108 CPD4_RAT CYT...	665	0.0
CRA 18000005107537 /altid=gi 2575863 /def=dbj BAA23125.1  (AB00...	665	0.0

#### EST:

Sequences producing significant alignments:	Score (bits)	E Value
gi 9872134 /dataset=dbest /taxon=960...	775	0.0
gi 6144331 /dataset=dbest /taxon=9606 ...	648	0.0
gi 6703894 /dataset=dbest /taxon=9606 ...	648	0.0

#### EXPRESSION INFORMATION FOR MODULATORY USE:

gi|9872134 /liver  
gi|6144331 /kidney  
gi|6703894 /lung

#### Tissue Expression:

Whole Liver

FIGURE 1

```

1 MGLEALVPLA VIVAIFLLLV DLMHRRQRWA ARYSPGPLPL PGLGNLLHVD
51 FQNTPYCFDQ LRRRFGDVFS LQLAWTPVVV LNLAAVREA LVTHGEDTAD
101 RPPVPITQIL GFGPRSQGRP FRPNGLLDKA VSNVIASLTC GRRFEYDDPR
151 FLRLDLAQE GLKEESGFLR EVLNAVPLL HIPALAGKVL RFQKAFLTQL
201 DELLTEHRMT WDPAQPPRDL TEAFLAEMEK AKGNPESSFN DENLRIVVAD
251 LFSAGMVTTS TTLAWGLLLM ILHPDVQRRV QQEIDDVIGQ VRRPEMGDQA
301 HMPYTTAVIH EVQRFGDIVP LGVTHMTSRD IEVQGFRIPK GTTLITNLSS
351 VLKDEAVWEK PFRFHEPEFL DAQGHFVKPE AFLPFSAGRR ACLGEPLARM
401 ELFLFFTSLL QHFSFSVPTG QPRPSHGVF AFLVTPSPYE LCAVPR
(SEQ ID NO: 2)

```

**FEATURES:**

**Functional domains and key regions:**

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

347-350 NLSS

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

327-329 TSR

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 5

1	93-96	THGE
2	198-201	TQLD
3	238-241	SFND
4	327-330	TSRD
5	437-440	SPYE

[4] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 2

1	233-238	GNPESS
2	255-260	GMVTTS

[5] PDOC00009 PS00009 AMIDATION  
Amidation site

Number of matches: 2

1	140-143	CGRR
2	387-390	AGRR

[6] PDOC00081 PS00086 CYTOCHROME\_P450  
Cytochrome P450 cysteine heme-iron ligand signature

385-394 FSAGRRACLG

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	3	23	1.877	Certain
2	68	88	1.096	Certain
3	171	191	0.668	Putative
4	252	272	1.914	Certain
5	400	420	1.402	Certain
6	425	445	0.833	Putative

BLAST Alignment to Top Hit:

>CRA|18000004889269 /altid=gi|181304 /def=gb|AAA53500.1| (M33388)  
cytochrome P450 IID6 [Homo sapiens] /org=Homo sapiens  
/taxon=9606 /dataset=nraa /length=497  
Length = 497

Score = 884 bits (2259), Expect = 0.0  
Identities = 444/497 (89%), Positives = 445/497 (89%), Gaps = 51/497 (10%)

Query: 1 MGLEALVPLAVIVAIFLLLVDMHRRQRWAARYSPGPLPLPGLGNLLHVDFQNTPYCFDQ 60  
MGLEALVPLAVIVAIFLLLVDMHRRQRWAARY PGPLPLPGLGNLLHVDFQNTPYCFDQ  
Sbjct: 1 MGLEALVPLAVIVAIFLLLVDMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQ 60

Query: 61 LRRRFGDVFSQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQG-- 118  
LRRRFGDVFSQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQG  
Sbjct: 61 LRRRFGDVFSQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF 120

Query: 119 -----RPFPRNGLLDK 129  
RPFPRNGLLDK  
Sbjct: 121 LARYGPAWREQRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFPRNGLLDK 180

Query: 130 AVSNVIASLTCGRREFYDDPRFLRLDLAQEGLKEESGFLREVLNAVVPVLLHIPALAGKV 189  
AVSNVIASLTCGRREFYDDPRFLRLDLAQEGLKEESGFLREVLNAVVPVLLHIPALAGKV  
Sbjct: 181 AVSNVIASLTCGRREFYDDPRFLRLDLAQEGLKEESGFLREVLNAVVPVLLHIPALAGKV 240

Query: 190 LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEAKAGNPESFNDENLRIVVA 249  
LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEAKAGNPESFNDENLRIVVA  
Sbjct: 241 LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEAKAGNPESFNDENLRIVVA 300

Query: 250 DLFSAGMVTSTTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI 309  
DLFSAGMVTSTTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI  
Sbjct: 301 DLFSAGMVTSTTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI 360

Query: 310 HEVQRFQGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFHFHPEHF 369  
HEVQRFQGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFHFHPEHF  
Sbjct: 361 HEVQRFQGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFHFHPEHF 420

Query: 370 LDAQGHFVKPEAFLPFSAGRRACLGEP LARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV 429  
LDAQGHFVKPEAFLPFSAGRRACLGEP LARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV  
Sbjct: 421 LDAQGHFVKPEAFLPFSAGRRACLGEP LARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV 480

Query: 430 FAFLVTPSPYELCAVPR 446  
FAFLV+PSPYELCAVPR  
Sbjct: 481 FAFLVSPSPYELCAVPR 497 (SEQ ID NO: 4)

Hammer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00067	Cytochrome P450	516.7	1.7e-151	2

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00067	1/2	35	113 ..	1	92 [.	78.1	2.7e-21
PF00067	2/2	117	443 ..	150	497 .]	442.7	3.3e-129

1 AGCCTTACAA AGTGTGCGGA TTACCTGCGT GAGCCACCGG GTCCGGCCTC  
51 TTTATGTCTT ACTGTAAGT CTGTCTTGAA AAGTACTTAT TATTTTGTAG  
101 TGGTTTCATCA TTTAGTCTAA TTTAAATAAG AGTAGTTTAC ACACCACAAT  
151 TACAGTATTA TAATACTCTG TTTTCTGTG TGCTTACTAT TACCAGTGAG  
201 TTTTGTACCT TTAGATGATT TCTTCTTGCT CATTAAATATC CTTTTTTTTT  
251 TCAGATTGAA AAACCTCCCT TAGCATTCTT TGTGGGATAT AGGTCTGGTG  
301 TTGATGAAAT CTCGAGCTT TTGTTTGTCT GGAAGGTCT TATTTCTCTC  
351 TTCCTGTTGG AAGGATATTT TTGCCAGATA CGTTATTCTA GGCTAAAAGT  
401 TTTTTTCTCT TCAGCACTTT AAATATGTCA TGCCACTCCC CCCTGGCCTG  
451 TAAGGTTTCC ACTGAAAAGG TGGCTGCCCC ATGTCATGTA TTGGAGCTCT  
501 ACTGCATGTT ATTTGTTTCT TTTCTCTTGC TGCTTTTAGG ATCCTTTCTT  
551 TATCCTTTGAC CTTTCGGAGT TTAATTATCA GATGCCTTGA GGTCTCTTTC  
601 TTTGGGTTAA ATCTGCTTGG TGTCTATAA ACTTCTTGTA CAAAAAATCA  
651 GCCAGGCATG GTGGTGGGCA CCTGTAATCC CAGCTACTTG GGAGGCTGAG  
701 GCAGGAGAAT CGCTTGAACC CTGGAGGTGG AGGTGTCAGT GAGCCGAGAT  
751 CGCATCTTAC CACTCCCACC TGGGCGACAG AGCAAAACTC CGTCTCAAAA  
801 AAAAAATTAT TTGGGCTCGG TGGTGCCTGT AGTCCCAGCT ACTTGGGAGG  
851 CAGGAGGTCC ACTTGATGTT GAGATTGCAG TGAGCCATGA TCCTGCCACT  
901 GCACTCCGGC CCGGGCAACA GAGTGAGACC CTGTCTAAAG AAAAAATAAA  
951 AATAAAAAAG CAACATATCC TAAATAAAGG ATCCTCCATA ATGTTTCCAC  
1001 CAGATTTCTA ATCAGAAACA TGGAGGCCAG GAAGCAGTGG AGAATGACGA  
1051 CCCTCAGGCA GCCCTGGAGG ATGCTGTAC AGGCTGGGGC AAGGGCCTTC  
1101 AGGCTACCAA CTGGGAGCTC TGGGAACAGC CCTGTTGCAA ACAGGAAGTC  
1151 ATGGCCCGGC CAGAGCCAG AATGTGGGCT GAGCTGGGAT CCATGTGACA  
1201 GCTTTGAGGC TCACCGGGAG CAGCCTCTGG ACAGGAGAGG TCCCATCCAG  
1251 GAAACCTCGG GCATGGCTGG GAAGTGGGGT ACTTGGTGCC GGGTCTGTAT  
1301 GTGTGTGTGA CTGGTGTGTG TGAGAGAGAA TGTGTGCCCT GAGTGTCACT  
1351 GTGAGTCTGT GTATGTGTGA ATATTGTCTT TGTGTGGGTG ATTTTCTGCA  
1401 TGTGTAATCG TGTCCCTGCA AGTGTGAACA AGTGGACAAG TGTCTGGGAG  
1451 TGGACAAGAG ATCTGTGCAC CATCAGGTGT GTGCATAGCG TCTGTGCATG  
1501 TCAAGAGTGC AAGGTGAAGT GAAGGGACCA GGCCCATGAT GCCACTCATC  
1551 ATCAGAGGCT CTAAGGCCCC AGGTAAGTGC CAGTGACAGA TAAGGGTGCT  
1601 GAAGGTCACT CTGGAGTGGG CAGGTGGGGG TAGGGAAAGG GCAAGGTCAT  
1651 GTTCTGGAGG AGGGGTGTG ACTACATTAG GGTGTATGAG CCTAGCTGGG  
1701 AGGTGGATGG CCGGGTCCAC TGAGACCTTG GTTATCCCAAG AACCTGTGT  
1751 GGGCTTGGGG AGCTTGGAGT GGGGAGAGGG GGTGACTTCT CCGACAGGC  
1801 CTTTCTACCA CCCTACCCTG GGTAAAGGCC TGGAGCAGGA AGCAGCGGCA  
1851 AGGACCTCTG GAGCAGCCCA TACCTGCCCT GGCTGACTC TGCCACTGGC  
1901 AGCACAGTCA ACACAGCAGG TCACTCACA GCAGAGGGCG AAGGCCATCA  
1951 TGACCTCCCT TTATAAGGA AGGGTCACGC GCTCGGTGTG CCGAGAGTGT  
2001 CCTGCCTGGT CCTCTGTGCC TGGTGGGGTG GGGGTGCCAG GTGTGTCCAG  
2051 AGGAGCCAG TTGGTAGTGA GGCAGCCATG GGCTAGAAG CACTGTGCTC  
2101 CCTGGCCATG ATAGTGGCCA TCTTCTGCT CCTGGTGGAC CTGATGCACC  
2151 GGCACCAACG CTGGGCTGCA CGCTACCCGC CAGGTCCCCT GCCACTGCC  
2201 GGGCTGGGCA ACCTTGCTGC ATGTGGACTT CCAGAACACA CCATACTGCT  
2251 TCGACAGGTT GAGGAGGAG GTCTTGGAGG GCGGCAGAGG TCCTGAGGAT  
2301 GCGCCACCAC CAGCAACAT GGGTGGTGGG TTAACACACA GGCTGGATCA  
2351 GAAGCGAGGC TGAGAAGGGG AAGCAGGTTT GGGGACGTTT CCTGGGGAAG  
2401 GACATTTATA CATGGCATGA AGGACTGGAT TTTCCAAAGG CCAAGGAAGA  
2451 GTAGGGCAAG GGCCTGGAGG TGGAGCTGGA CTGGCAGTG GGCATGCAAG  
2501 CCCATTGGGC AACATATGTT ATGGAGTACA AAGTCCCTTC TGCTGACACC  
2551 AGAAGGAAAG GCCTTGGGAA TGGAAGATGA GTTAGTCTCT AGTGCCGTTT  
2601 AAATCACGAA ATCGAGGATG AAGGGGGTGC AGTGACCCGG TTCAAACCTT  
2651 TTGCACTGTG GGTCTCTGGG CCTCACTGCT CACCGGCATG GACCATCATC  
2701 TGGGAATGGG ATGCTAACTG GGGCCTCTCG GCAATTTTGG TGACTCTTGC  
2751 AAGGTACATC CTGGGTGACG CATCCAAACT GAGTTCCTCC ATCAGAGAAG  
2801 GTGTGACCCC CACCCCTGCC CCACGATCAG GAGGCTGGGT CTCCTCCTTC  
2851 CACCTGCTCA CTCCTGGTAG CCCCGGGGGT CGTCCAAGGT TCAAATAGGA  
2901 CTAGGACCTG TAGTCTGGGG TGATCCTGGC TTGACAAGAG GCCCTGACCC  
2951 TCCCTCTGCA GTTGGCGGCG CGCTTGGGGG ACGTGTTCAG CCTGCAGCTG  
3001 GCCTGGACGC CGGTGGTCTG GCTCAATGGG CTGGCGGGCG TGCGCGAGGC  
3051 GATGGTGACC CGCGGCGAGG ACACGGCCGA CCGCCCGCCT GCGCCCATCT  
3101 ACCAGTCTCT GGGCTTCCGG CCGCTTCCC AAGGCAAGCG GCGGTGGGGG  
3151 ACAGAGACCG CGTTTCCGTG GGGCCCGGGT GGACAGTGAC CGTAGCCCAA  
3201 GCAGCGCCGA CAGGCGGTGG GTCTCTGGAC GTGAAACAGA GATAAAGGCC  
3251 AGCGAGTGGG CTGAGGACAG TGGGCCAGGA AACCACCTGC ACGGGGGAGG  
3301 TGCGAGTCTG TGGGCTGGGA GGGGGCGGGG CTAAGTCCCA GACCCGCCAG  
3351 AAGCCCGGTG GCGGAGGCTG ATGCGTCGAA GTGGCGGTGG CGGGGACCGC  
3401 CCTATGCTG CGGGCTCAGT GTGGGCGGGA CCGGCGGGAT CTTCTTGTAG  
3451 TGGAAAGGTG GTCAGGCTGG GCAGAGACGA GGTGGGGCCA AACCCCGCCC  
3501 CAGGCAGGGG AGCAATGTGG GTGAGCAAAG AGTGGGCCCT GTGCCAGCT

```

3551 GGACCGGGCT AGGGACTGCG GGAGACCTTG TGGAGCGCCA GGGTTGGAGT
3601 GGGTGGCGGA GGGTGGGGCC AAGGCCTTCA TGGCAACGCC CACGTGTCCG
3651 TCCCGCCCCC AGGGGTGATC CTGTCGCGCT ATGGGCCCGC GTGGCGCGAG
3701 CAGAGCGCCT TCTCCGTGTC CACCTTGCGC AACTTGGGCC TGGGCAAGAA
3751 GTCGCTGGAG CAGTGGGTGA CCGAGGAGGC CGCCTGCCTT TGTGCCGCTT
3801 CGCCGACCAA GCCGGTGGGT GATGGGCAGA AGGGCACAAA CGCGGAACTG
3851 GGAAGCGCGG GGACGGAGAA GGCAACCCCT TACCCGCATC TCCCCACCCC
3901 CAGGACGCCC CTTTCGCCCC AACGGCCTCT TGGACAAAGC CGTGAGCAAC
3951 GTGATCGCCT CCCTCACCTG CGGGCGCCGC TTCGAGTACG ACGACCCTCG
4001 CTTCTCAGG CTGCTGGACC TAGCTCAGGA GGGACTGAAG GAGGAGTCGG
4051 GCTTTCTGCG CGAGGTGCGG AGCGAGAGAC CGAGGAGTCT CTGCAGGGCG
4101 AGCTCCTGAG AGGTGCGGG GCTGGACTGG GGCCCTCCGA GGGCAGGATT
4151 TGCATAGATG GGTTTGGGAA AGGACATTCC AGGAGACCCC ACTGTAAGAA
4201 GGGCCTGGAG GAGGAGGGGA CATCTCAGAC ATGGTCGTGG GAGAGGTGTG
4251 CCCGGGTGAG GGGGCACCAG GAGAGGCCAA GGAATCTGTA CCCCCGTCCA
4301 CGTTGGAGAT TTCGATTTTA GGTTCCTCCT CTGGGCAAGG AGAGAGGGTG
4351 GAGGCTGGCA CTTGGGGAGG GACTTGGTGA GGTGAGTGGT AAGGACAGGC
4401 AGGCCTTGGG TCTACCTGGA GATGGCTGGG GCCTGAGACT TGTCCAGGTG
4451 AACGCAGAGC ACAGGAGGGA TTGAGACCCC GTTCTGTCTG GTGTAGGTGC
4501 TGAATCTGTG CCCCCTCCTC CTGCACATCC CAGCGCTGGC TGGCAAGGTC
4551 CTACGCTTCC AAAAGGCTTT CCTGACCCAG CTGGATGAGC TGCTAACTGA
4601 GCACAGGATG ACCTGGGACC CAGCCCAGCC ACCCCGAGAC CTGACTGAGG
4651 CCTTCTTGGC AAAGAAGGAG AAGGTGAGAG TGGCTGCCAC GGTGGGGGGC
4701 AAGGTGAGTG GGTGGAACGT CCGAGGAGGA ATGAGGGGAG GCTGGGCAAA
4751 AGGTGAGACC AGTGCATCAC CCGCGAGGCC GCATCTGGGC TGACAGGTGC
4801 AGAATGAGAG GTCAATTTGG GGCTACCCCG TTCTATCCCC TGAGTATCCT
4851 CTCGGCCCTG CTCAGGCCAA GGGGAGCCCT GAGAGCAGCT TCAATGATGA
4901 GAACCTGCGC ATAGTGGTGG GTAACCTGTT CCTTGCCGGG ATGGTGACCA
4951 CCTCGACCAC GCTGGCCTGG GGCCCTCCTG TCATGATCCT ACACCTGGAT
5001 GTGCAGCGTG AGCCAGCTG GGGCCCAAGG CAGGGACTGA GGGAGGAAGG
5051 GTACAGCTGG GGGCCCTTGG GCTTAGCTGG GACACCCGGG GCTTCCAGCA
5101 CAGGCGTGGC CAGGCTCCTG TAAGCCTAAC TTCCTCAAC ACAGGAGGAA
5151 GGAGAGTGTC CCCTGGGTGC TGACCCATTG TGGGGACGCA TGTCTGTCCA
5201 GTCCGTGTCC AACAGGAGAT CGACGACGTG ATAGGGCAGG TGCGGCGACC
5251 AGAGATGGGT GACCAGGCTC ACATGCCCTA CACCACTGCC GTGATTCACG
5301 AGGTGAGCGC CTTTGGGGAC ATCATCCCCC TGAGTGTGAC CCATATGACA
5351 TCCCGTGACA TCGAAGTACA GGGCTTCCGC ATCCCTAAGG TAGGCCTGGC
5401 GCCCTCCTCA CCCAGCTCA GCACAGCAC CTGGTGATAG CCCCAGCATG
5451 GCTACTGCCA GGTGGGCCCC CTCTAGGAAC CCTGGCCACC TAGTCCTCAA
5501 TGCCACCACA CTGACTGTCC CCACTTGGGT GGGGGGTCCA GAGTATAGGC
5551 AGGGCTGGCC TGTCCATCCA GAGCCCCCGT CTAGTGGGGA GACAAACCAG
5601 GACCTGCCAG AATGTTGGAG GACCCAGCGC CTGCAGGGAG AGGGGGCAGT
5651 GTGGGTGCCT CTGAGAGGTG TGAATGCGCC CTGCTGTGGG GTCGGAGAGG
5701 GTACTGTGAG GTTCTCGGG CGCAGGACTA GTTGACAGAG TCCAGCTGTG
5751 TGCCAGGCAG TGTGTGTCCC CCGTGTGTTT GGTGGCAGGG GTCCCAGCAT
5801 CCTAGAGTCC AGTCCCCACT CTCACCTGCG ATCTCCTGCC CAGGGAACGA
5851 CACTCATCAC CAACCTGTCA TCGGTGCTGA AGGATGAGGC CGTCTGGGAG
5901 AAGCCCTTCC GCTTCCACCC CGAACACTTC CTGGATGCC AGGGCCACTT
5951 TGTGAAGCCG GAGGCCCTTC TGCCCTTCTC AGCAGGTGCC TGTGGGGAGC
6001 CCGGCTCCCT GTCCCTTCC GTGGAGTCTT GCAGGGGTAT CACCCAGGAG
6051 CCAGGCTCAC TGACGCCCCC CCCTCCCCA CAGGCCCGCC TGCATGCCTC
6101 GGGGAGCCCC TGGCCCGCAT GGAGCTCTTC CTCTTCTTCA CCTCCCTGCT
6151 GCAGCACTTC AGCTTCTCCG TGGCCGCGCG ACAGCCCCCG CCCAGCCACT
6201 CTCGTGTCGT CAGCTTTCTG GTGACCCCAT CCCCCTACGA GCTTTGTGCT
6251 GTGCCCCGCT AGAATGGGGT ACCTAGTCCC CAGCCTGCTC CCTAGCCAGA
6301 GGCTCTAATG TACAATAAAG CAATGTGGTA GTTCCAATT GGGTCCCCTG
6351 CTCACGCCCT CGTTGGGATC ATCCTCCTCA GGGCAACCCC ACCCCTGCCT
6401 CATTCCTGCT TACCCACCG CCTGGCCGCA TTTGAGACGG GTACGTTGAG
6451 GCTGAGCAGA TGTGAGTTAC CCTTGCCCAT AATCCCATGT CCCCCTACTG
6501 CCAAACCTG ACTGCCCAGA TTGGTGACAA GGAATACATT GTCCTGGCAT
6551 GTGGGGGAGG GGCCAGAATG GGCTGACTAG AGGTGTCAGT CAGCCCTGGA
6601 TGTGGTGGAG AGGGCAGGAC TCAGCCTGGA GGCCCATATT TCAGGCCTAA
6651 CTCAGCCAC CCCACATCAG GGACAGCAGT CCGCCAGCA CCATCACAAAC
6701 AGTCACCTCC CTTATATAT GACACCCCAA AATGGAAGAC AAATCATGTC
6751 AGGGAGCTAT ATCCAGGGC TACCTCCAG GGCTCAGTCG GCAGGTGCCA
6801 GAACATTCCC TGGGAAGGCC CCAGGAAAAC CCAGGACCGA GCCACCGCCC
6851 TCAGCCTGTC ACCTGTGTG CAAAATTGGT GGGTTCTTGG TCTCACTGAC
6901 TTCAAGAATG AAGCCGTGGA CCCTCACGGT GAGTGTTACA GTTCTTAAAG
6951 ATGGTGTGTT CAGAGTTTGT TCCTTCTGAT GTTAAGACGT GTTCAGAGTT
7001 TCTTCTTCTT GGTGGGTGCG TGGTCTTGCT GGCTTCAGGA GTGAAGCTGC
7051 AGACCTTCAC AGTGAGTGT ACGGCTCTTA AGGCTGCACG TACGGAGTTG

```

FIGURE 3, page 2 of 5

```

7101 TTCATTCTTC CTGGTGGGTT TGTGGTCTCA CTGGCCTCAG GAGTGAAACT
7151 GCAGTCCCTC CAGTGTTACA ACTCATAAAG GCAGTGTGGA CCCAATGAGG
7201 GAGCAGCAGC AGCAAGACTT ACTGCAACA GCAAAAGAAT GATGGCAACC
7251 AGGTTGCCCG TGCTACTTCA GGCAGCCTGC TTTTATTCCC TTATCTGACC
7301 CCCACCCACA TCCTGCTGAT TGGCCCATTT TACAGACAGT GGATTGGTCC
7351 ACTTACAGAG AGCTGATTGG TGCATTTACA ATCCCTGAGC TAGACACAGA
7401 GTAGTGATTG GTATATTTAC AAACCTTGAG CTAGACACAG AGTGCTGAAT
7451 GGTGTATTTC CAATCCCTTA GCTAGACATA AAGGTTGTCC CAGTCCCCAC
7501 TAGATTAGCT AGATAGAGTA GACAGAGAGC ACTGATTGGT GCGTTTACAA
7551 ACCTTGAGTT AGACACAGGG TGCTGACTGG TGTGTTTACA AACCTTGAGC
7601 TAGACACAGA GTGCTGATTG GTGATTTTAC AATCTTTTAG CTAGAAATAA
7651 AGGTTGCCCG AGTCCCCACC AGATTAGCTA GATAGAGTGC TAATTGGTGC
7701 ATGCACGAAC CCGGAGCTAG ACACAGAGTG CTGATTGGTG CATATACAAT
7751 CCTCTGGCTA GACATAAAAG TTCTCCAAGT CCCACCTGA CTCAGGAGCC
7801 CAGCCAGCTT CGCCTAGTGG ATCCTATGCC AGGGCCACAG GCAGAGCTGC
7851 CTGCTAGTCC CACACCGGGC ACCTGTACTC CTCAGCCCTT GGGCAGTGGA
7901 CGGGACCAGG TGCCGTGGAG CAGTGGGAGG CACCCATCCG GGAGGCTCGG
7951 GCCTGCGAGG GAGCCACCG TAGGGAGGCT TGGGCATGGC AGGCTGCAAG
8001 TCCTGAGCCC TGCCCCGCGG GGAGGTGACT GAGGCCTGGC GACAATTCAA
8051 GTGTGGTGAG AGCCCGCAGG CCAGCAGTAC TGGGGGACCC GGTGCCCCCT
8101 CTGCAGCTGC TGGCCAGGT GCTAAGCCCC TCACTGCCTG GGGCCAGAGG
8151 CACCAGCCGG CCGTCCGAG TGCAGGGCCC GCTGAGCCCC TGCCCCACCA
8201 GAACTGGTGC TGGCCCGCGA GCAACCCAGG TTCCCGCACA CGCCTCTCCC
8251 TCATATCCCT CCGCAAGCA GACGGAGCCG GCTCCAGCCT CCACCAGTCC
8301 AGAGAGGGGC TCCCACAGTG CAGCGCTGGG CTGAACAAGG TCCTACGCTT
8351 CCAAAGGCTT TTCTGACCC AGCTGGATGA GCTGCTAACT GAGCACAGGA
8401 TGACCTGGGA CCCAGCCCAG CCCCCCGGAG ACCTGACTGA GGCCTTTCTT
8451 GGCAGAGATG GAGAAGGTGA GAGTGGCTGC CACGGTGGGG GGCAAGGGTG
8501 GTGGGTGTAG CGTCCCAGGA GGAATGAGGG GAGGCTGGGC AAAAGGTTGG
8551 ACCAGTGCAT CACCCGGCGA GCCGCATCTG GGCTGACAGG TGCAGAATTG
8601 GAGGTCATTT GGGGGCTACC CCGTTCTGTC CCGAGTATGC TCTCGGCCCT
8651 GCTCAGGCGA AGGGGAACCC TGAGAGCAGC TTCAATGATG AGAACCTGCG
8701 CATAGTGGTG GGTGACCTGT TCTTGCCGGG GATGGTGACC ACCTCGACCA
8751 CGCTGGCCTG GGGCTCCTG CTCATGATCC TACATCCGGA TGTGCAGCGT
8801 GAGCCCATCT GGGAAACAGT GCAGGGGCGG AGGGAGGAAG GGTACAGGCG
8851 GGGGCCCATG AACTTTGCTG GGACACCCGG GGCTCCAAGC ACAGGCTTGA
8901 CCAGGATCCT GTAAGCCTGA CCTCCTCCAA CATAGGAGGC AAGAAGGAGT
8951 GTCAGGGCCG GACCCCTGG GTGCTGACCC ATTGTGGGGA CGCATGTCTG
9001 TCCAGGCCGT GTCCAACAGG AGATCGACGA CGTGATAGGG CAGGTGCGGC
9051 GACCAGAGAT GGGTGACCAG GCTCACATGC CCTACACCAC TGCCGTGATT
9101 CATGAGGTGC AGCGCTTTGG GGACATCGTC CCCCTGGGTG TGACCCATAT
9151 GACATCCCGT GACATTCGAA GTACAGGGCT TCCGCATCCC TAAGGTAGGC
9201 CTGGCGCCNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNCT GCCCAGGGAA CGACACTCAT CACCAACCTG TCATCGGTGC
9351 TGAAGGATGA GGCCGTCTGG GAGAAGCCCT TCCGCTTCCA CCCCGAACAC
9401 TTCCTGGATG CCCAGGGCCA CTTGTGTAAG CCGGAGGCC TCCCTGCCTTT
9451 CTCAGCAGGT GCCTGTGGGG AGCCCGGCTC CCTGTCCCTT TCCGTGGAGT
9501 CTTGCAGGGG TATACCCAG GAGCCAGGCT CACTGACGCC CCTCCCTTCC
9551 CCACAGGCCG CCGTGCATGC CTCGGGGAGC CCCTGGCCCG CATGGAGCTC
9601 TTCCTCTTCT TCACCTCCCT GCTGCAGCAC TTCAGTTCT CGGTGCCAC
9651 TGGACAGCCC CGGCCAGCC ACCATGGTGT CTTTGCTTTC CTGGTGAGCC
9701 CATCCCCCTA TGAGCTTTGT GCTGTGCCCG GCTAGAATGG GGTACCTAGT
9751 CCCCAGCCTG CTCCTAGCC AGAGGCTCTA ATGTACAATA AAGCAATGTG
9801 GTAGTTCCAA CTCGGGTCCC CTGCTCACGC CCTCGTTGGG ATCATCTCTC
9851 TCAGGGCAAC CCCACCCCTG CCTCATTCCT GCTTACCCCA CCGCCTGGCC
9901 GCATTTGAGA CAGGGGTACG TTGAGGCTGA GCAGATGTCA GTTACCCTTG
9951 CCCATAATCC CATGTCCCCC ACTGACCCAA CTCTGACTGC CCAGATTGGT
10001 GACAAGGACT ACATTGTCTT GGCATGTGGG GAAGGGGCCA GAATGGGCTG
10051 ACTAGAGGTG TCAGTCAGCC CTGGATGTGG TGGAGAGGGC AGGACTCAGC
10101 CTGGAGGCCC ATATTTTACG CCTAACTCAG CCCACCCAC ATCAGGGACA
10151 GCAGTCCCTG CAGCACCATC ACAACAGTCA CCTCCCTTCA TATATGACAC
10201 CCCAAAACGG AAGACAAATC ATGGCGTCAG GGAGCTATAT GCCAGGGCTA
10251 CCTACCTCCC AGGGCTCAGT CGGCAGGT

```

(SEQ ID NO: 3)  
**FEATURES:**  
 Start.....2078  
 Exon: 2078-2258  
 Intron: 2259-2961  
 Exon: 2962-3133

Intron: 3134-3903  
 Exon: 3904-4064  
 Intron: 4065-4496  
 Exon: 4497-4673  
 Intron: 4674-4865  
 Exon: 4866-5007  
 Intron: 5008-5201  
 Exon: 5202-5389  
 Intron: 5390-5843  
 Exon: 5844-5985  
 Intron: 5986-9556  
 Exon: 9557-9732  
 Stop 9733

SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
3101	C	T A	Exon	107	T	T T
3439	A	G	Intron			
4908	C	T	Exon	245	P	L
5627	G	A	Intron			
6733	T	C	Intron			
7788	-	C T	Intron			
7867	G	A	Intron			
7948	C	T	Intron			

Context:

DNA  
Position

3101 GTGTGACCCCCACCCCTGCCCCACGATCAGGAGGCTGGGTCTCCTCCTTCCACCTGCTCA  
 CTCTGTGTAGCCCCGGGGGTCTGCCAAGGTTCAAATAGGACTAGGACCTGTAGTCTGGGG  
 TGATCCTGGCTTGACAAGAGGCCCTGACCCCTCCCTCTGCAGTTGCGGCGCCGCTTCGGGG  
 ACGTGTTTCAGCCTGCAGCTGGCCTGGACGCCGTGGTCTGCTCAATGGGCTGGCGGCCG  
 TGCGCGAGGCGATGGTGACCCGCGGCGAGGACACGCCCGACCGCCCGCTGCGCCCATCT  
 [C, T, A]  
 CCAGGTCTGGGCTTCGGGCCCGCTTCCCAAGGCAAGCGGCGGTGGGGGACAGAGACCGC  
 GTTTCCTGGGCCCCGGTGGACAGTGACCGTAGCCCAAGCAGCGCCGACAGGGCGTGGG  
 GTCCTGGACGTGAAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAA  
 ACCACCTGCACGGGGGAGGTGCGAGTCTGTGGCTGGGAGGGGCGGGGCTACTGCCAG  
 ACCCGCCAGAAGCCCGTGGGCGAGGCTGATGCGTCAAGTGGCGGTGGCGGGGACCGCG

3439 CGGCGGTGGGGGACAGAGACCGGCTTCCGTGGGCCCCGGGTGGACAGTGACCGTAGCCC  
 AAGCAGCGCCGACAGGGCGTGGGTCTTGACGTGAAACAGAGATAAAGGCCAGCGAGTG  
 GGCTGAGGACAGTGGGCCAGGAAACACCTGCACGGGGGAGGTGCGAGTCTGTGGGCTGG  
 GAGGGGGCGGGGCTACTGCCAGACCCGCCAGAAGCCCGTGGGCGAGGCTGATGCGTCG  
 AAGTGGCGGTGGCGGGGACCGCGCCTATGCTGCGGGCTCAGTGTGGGCGGGACGGCGGG  
 [A, G]  
 TCTTCTTGTAGTGAAAGGTGGTCAGGGTGGGCAGAGACGAGGTGGGGCCAAACCCCGCC  
 CCAGGCAGGGGAGCAATGTGGGTGAGCAAAGAGTGGGCCCTGTGCCAGCTGGACCGGGC  
 TAGGGACTGCGGGAGACCTTGTGGAGCGCCAGGGTTGGAGTGGGTGGCGGAGGGTGGGGC  
 CAAGGCCCTCATGGCAACGCCACGTGTCCGTCCCGCCCCAGGGGTGATCTGTCTGCGC  
 TATGGGCCCGCTGGCGGAGAGAGCGCTTCTCCGTGTCCACCTTGCGCAACTTGGGC

4908 ATGACCTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGGCCTTCCTGGCAAAGAAG  
 GAGAAGGTGAGAGTGGCTGCCACGGTGGGGGCAAGGTTGGTGGTTGAACGTCCCAGGA  
 GGAATGAGGGGAGGCTGGGCAAAAGGTTGGACAGTGCATCACCCGCGAGCCGCTCTG  
 GGCTGACAGGTGCAGAAATTGGAGGTCATTTGGGGGCTACCCCGTTCTATCCCTGAGTAT  
 CCTCTCGGCCCTGCTCAGGCCAAGGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACTG  
 [C, T]  
 GCATAGTGGTGGGTAACCTGTTCTTGCCTGGGATGGTGACCACCTCGACCAGCTGGCCT  
 GGGGCCCTCTGCTCATGATCCTACACCTGGATGTGCAGCGTGAGCCAGCTGGGGCCCAA  
 GGCAGGACTGAGGGAGGAAGGTTACAGCTGGGGGCCCTGGGCTTAGCTGGGACACCCG  
 GGGCTTCCAGCACAGGCGTGGCCAGGCTCCTGTAAGCCTAACTTCTTCAACACAGGAGG  
 AAGGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGGACGCATGTCTGTCCAGTCCGTGT

5627 CCCCTGAGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGCTTCCGCATCCCT



AAGGTAGGCCTGGCGCCTCCTCAGCCAGCTCAGCACCAGCACCTGGTGATAGCCCCAG  
CATGGTACTGCCAGGTGGGCCCCACTCTAGGAACCTTGGCCACCTAGTCTCAATGCCAC  
CACACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGCAGGGCTGGCCTGTCCA  
TCCAGAGCCCCGTCTAGTGGGAGACAAACCAGGACCTGCCAGAATGTTGGAGGACCCA  
[G, A]  
CGCCTGCAGGGAGAGGGGGCAGTGTGGGTGCCTCTGAGAGGTGTGACTGCGCCCTGCTGT  
GGGGTCGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCT  
GTGTGCCAGGCAGTGTGTGTCCCCGTGTGTTGGTGGCAGGGGTCCCAGCATCCTAGAG  
TCCAGTCCCCACTCTCACCCTGCATCTCCTGCCCAGGGAACGACACTCATCACCAACCTG  
TCATCGGTGCTGAAGGATGAGGCCGTCTGGGAGAAGCCCTTCCGCTTCCACCCGAACAC

6733 TGAGACGGGTACGTTGAGGCTGAGCAGATGTCAGTTACCCTTGCCCATATCCCATGTCC  
CCCAGTACCCAACTCTGACTGCCCAGATTGGTGACAAGGACTACATTGTCTGGCATGT  
GGGGAAGGGGGCAGAAATGGGCTGACTAGAGGTGTCACTCAGCCCTGGATGTGGTGAGAG  
GGCAGGACTCAGCCTGGAGGGCCATATTTAGGCCCTAACTCAGCCACCCACATCAGGG  
ACAGCAGTCTGCCAGCACCATCACACAGTCACCTCCCTTCATATATGACACCCCAAAA  
[T, C]  
GGAAGACAAATCATGTGAGGGAGCTATATGCCAGGGCTACCTCCCAGGGCTCAGTCGGCA  
GGTGCCAGAACATTCCCTGGGAAGGGCCCCAGGAAACCCAGGACCGAGCCACCGCCCTCA  
GCCTGTCACTTGTGTCCAAATTTGGTGGGTCTTGGTCTCACTGACTTCAAGAATGAAG  
CCGTGGACCTCAGGTGAGTGTACAGTTCTTAAAGATGGTGTGTTTCAAGTTTGTTC  
TTCTGATGTTAAGACGTGTTTCAAGTTTCTTCTTCTGGTGGGTGCGTGGTCTTGTGCG

7788 TCCAGTCCCCACTAGATTAGCTAGATAGAGTAGACAGAGAGCACTGATTGGTGCCTTTA  
CAAACCTTGAGTTAGACACAGGGTGTGACTGGTGTGTTTACAAACCTTGAGCTAGACAC  
AGAGTGTCTGATTGGTGTATTTACAATCTTTTAGCTAGAAATAAAGGTTCCCAAGTCCCC  
ACCAGATTAGCTAGATAGAGTGTAAATTGGTGCATGCACGAACCCGGAGCTAGACACAGA  
GTGCTGATTGGTGCATATACAATCCTCTGGCTAGACATAAAAGTTCTCCAAGTCCCCACC  
[-, C, T]  
GACTCAGGAGCCCAGCCAGCTTCGCCTAGTGGATCCTATGCCAGGGCCACAGGCAGAGCT  
GCCTGCTAGTCCCACACCGGGCACCCTGTACTCTCAGCCCTTGGGCAGTGGACGGGACCA  
GGTGCCGTGGAGCAGTGGGAGGCACCCATCCGGGAGGCTCGGGCCTCGCAGGGAGCCAC  
CGTAGGGAGGCTTGGGCATGGCAGGCTGCAAGTCTGAGCCCTGCCCCGCGGGAGGTGA  
CTGAGGCCTGGCGACAATTCAAGTGTGGTGAGCGCCGCGCAGGCCAGCAGTACTGGGGGAC

7867 AGGGTGTGACTGGTGTGTTTACAAACCTTGAGCTAGACACAGAGTGTGATTGGTGTAT  
TTACAATCTTTTAGCTAGAAATAAAGGTTCCCAAGTCCCCACCAGATTAGCTAGATAGA  
GTGCTAATTGGTGCATGCACGAACCCGGAGCTAGACACAGAGTGTGATTGGTGCATATA  
CAATCTCTGGCTAGACATAAAAGTTCTCCAAGTCCCCACCTGACTCAGGAGCCCAGCCA  
GCTTCGCCTAGTGGATCCTATGCCAGGGCCACAGGCAGAGTGCCTGCTAGTCCCACACC  
[G, A]  
GGCACCTGTACTCCTCAGCCCTTGGGCAGTGGACGGGACCAGGTGCGGTGGAGCAGTGGG  
AGGCACCCATCCGGGAGGCTCGGGCCTCGCAGGGAGCCACCGTAGGGAGGCTTGGGCAT  
GGCAGGCTGCAAGTCTGAGCCCTGCCCCGCGGGAGGTGACTGAGGCCTGGCGACAATT  
CAAGTGTGGTGGAGCGCGGCAGGCCAGCAGTACTGGGGGACCCGGTGGCCCTCTGCAGC  
TGCTGCCCCAGGTGCTAAGCCCTCACTGCCTGGGGCCAGAGGCACCCAGCGGCCGCTCC

7948 TAAAGTTCCCAAGTCCCCACCAGATTAGCTAGATAGAGTGTAAATTGGTGCATGCACG  
AACCCGGAGCTAGACACAGAGTGTGATTGGTGCATATACAATCCTCTGGCTAGACATAA  
AAGTTCTCAAGTCCCCACCTGACTCAGGAGCCCAGCCAGCTTCGCCTAGTGGATCCTAT  
GCCAGGGCCACAGGCAGAGTGCCTGCTAGTCCACACCGGGCACCCTGACTCCTCAGCC  
CTTGGGCAGTGGAGGGGACCAGGTGCCGTGGAGCAGTGGGAGGCACCCATCCGGGAGGCT  
[C, T]  
GGGCCTCGCAGGGAGCCACCGTAGGGAGGCTTGGGCATGGCAGGCTGCAAGTCTTGAGC  
CCTGCCCCGCGGGAGGTGACTGAGGCCTGGCGACAATTCAAGTGTGGTGGAGCGCCGCA  
GGCCAGCAGTACTGGGGGACCCGGTCCCCCTCTGCAGTGTGGCCAGGTGCTAAGCC  
CCTCACTGCCTGGGGCCAGAGGCACCCAGCGGCCGCTCCGAGTGCAGGGCCGCTGAGCC  
CCTGCCACCCAGAACTGGTGTGCCCCGCGAGCAACCCAGGTTCCCGCACACGCCTCTC

Chromosome mapping:  
Chromosome #22